UNITED STATES PATENT AND TRADEMARK OFFICE

m✓

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/628,464	07/29/2003	Jon Elliot Adler		4703

21967   7590   02/15/2007

HUNTON & WILLIAMS LLP
INTELLECTUAL PROPERTY DEPARTMENT
1900 K STREET, N.W.
SUITE 1200
WASHINGTON, DC 20006-1109

EXAMINER
HOWARD, ZACHARY C

ART UNIT	PAPER NUMBER
1646	

SHORTENED STATUTORY PERIOD OF RESPONSE	MAIL DATE	DELIVERY MODE
3 MONTHS	02/15/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

PTOL-90A (Rev. 10/06)

	Application No.	Applicant(s)
Office Action Summary	10/628,464	ADLER ET AL.
	Examiner	Art Unit	
	Zachary C. Howard	1646	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on <u>06 December 2006</u>.

2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) <u>93-117</u> is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) <u>93-117</u> is/are rejected.

7)☒ Claim(s) <u>93</u> is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☐ All   b)☐ Some * c)☐ None of:

      1.☐ Certified copies of the priority documents have been received.

      2.☐ Certified copies of the priority documents have been received in Application No. _____.

      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☒ Notice of References Cited (PTO-892)

2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3)☐ Information Disclosure Statement(s) (PTO/SB/08)
    Paper No(s)/Mail Date _____.

4)☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____.

5)☐ Notice of Informal Patent Application

6)☒ Other: *Sequence Alignments #1-4.*

DETAILED ACTION

Status of Application, Amendments and/or Claims

The amendment of 12/6/06 has been entered in full. Claims 68-92 are canceled. New claims 93-117 are added.

Claims 93-117 are under consideration in the instant application.

The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

In view of the new grounds of rejection, the finality of the previous Office Action is withdrawn.

Withdrawn Objections and/or Rejections

The following page numbers refer to the previous Office Action (9/6/2006).

All rejections of claims 68-92 are moot in view of Applicants' cancellation of these claims.

The rejection of claims 68-92 under 35 U.S.C. § 101 for lack of utility set forth at page 2-4 has not been applied to new claims 93-117 in view of Applicants' resubmission of the Affidavit pursuant to 37 CFR 1.132 and Figures 1-2 associated which has been fully considered by the Examiner and provides sufficient evidence of utility for the claimed nucleic acids.

The rejection of claims 68-92 under 35 U.S.C. 112, first paragraph for lack of enablement set forth at pg 4-7 has not been applied to new claims 93-117 in view of (1) Applicants' resubmission of the Affidavit pursuant to 37 CFR 1.132 and Figures 1-2 associated therewith which have been found to be sufficient; and (2) in view of the different claim limitations presented in new claims 93-117.

The rejection of claims 68-92 under 35 U.S.C. 112, first paragraph for lack of written description set forth at pg 7-10 has not been applied to new claims 93-117 in view of the different claim limitations presented in new claims 93-117.

The additional grounds for rejection of claim 92 under 35 U.S.C. 112, first paragraph for lack of written description set forth at pg 9-10 is also withdrawn on further

consideration by the Examiner; therefore, this rejection has not been applied to new claim 117. On further consideration by the Examiner, the human T2R sequences described in the specification provide sufficient written description in order to identify the T2R nucleic acid sequences from other species that are encompassed by claim 117.

The rejection of claims 68-73 under 35 U.S.C. § 102(a) as being anticipated by WO200257309-A1 (Miwa et al, published July 25, 2002) set forth at pg 12-14 has not been applied to new claims 93-117 in view of the combination of (1) Applicants' submission on 12/6/06 of a 131 Affidavit executed by all of the inventors except for Elliot Adler; and (2) Applicants' 131 submission on 9/3/2005 of a 131 Affidavit executed by Robin Teskin; this Affidavit has been considered effective in place of an Affidavit executed by Elliot Adler in view of the Petition under 37 C.F.R 1.47(a) granted 3/11/2004. It is noted that MPEP 715.04 [R-5] states, "The following parties may make an affidavit or declaration under 37 CFR 1.131...(c) If a petition under 37 CFR 1.47 was granted ... the affidavit or declaration may be signed by the 37 CFR 1.47 applicant or the legal representative, where appropriate." However, please see the new rejection under 35 U.S.C § 102(a) set forth below.

Note

At page 6 of the 12/6/06 response, Applicants state "Applicants respectfully note that the PTOL-326 incorrectly states that this Office Action is a final rejection. In fact the text of the Office Action correctly notes that the Office Action is a non-final rejection based on the RCE Request."

In response, the Examiner notes that the 9/6/06 Office Action was correctly indicated to be a final rejection both on the PTOL-326 and in the Office Action at pg 14, in accordance with MPEP § 706.07(b). Applicants' request for continued examination (RCE) under 37 CFR 1.114 was entered at the beginning of the 9/6/06 Office Action (see page 2). However, the conclusion of the Office Action (page 14) stated, "All claims are drawn to the same invention claimed in the application prior to the entry of the submission under 37 CFR 1.114 and could have been finally rejected on the grounds and art of record in the next Office action if they had been entered in the application

prior to entry under 37 CFR 1.114. Accordingly, **THIS ACTION IS MADE FINAL** even though it is a first action after the filing of a request for continued examination and the submission under 37 CFR 1.114. See MPEP § 706.07(b). Applicants are reminded of the extension of time policy as set forth in 37 CFR 1.136(a)."

However, in view of the new grounds of rejection set forth below, the finality of the previous Office Action is withdrawn.

Maintained Objections and/or Rejections
Claim Rejections - 35 USC § 112, 2nd paragraph

Claims 93-117 are rejected under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. This rejection was set forth for claims 68-92 at pg 11-12 of the 9/6/06 Office Action; these claims have been cancelled but the rejection is herewith applied to claims 93-117.

Claims 93-97 are indefinite because the metes and bounds of the phrase "contained in" are unclear. For example, in claim 93 it is unclear whether a "polypeptide contained in SEQ ID NO: 2" is limited to a polypeptide consisting of SEQ ID NO: 2, or whether it encompasses shorter polypeptides (fragments) that are "contained in" the longer sequence of SEQ ID NO: 2. For purposes of prosecution, claims 93-97 have been interpreted broadly to include polynucleotides comprising fragments of SEQ ID NO: 1, or polynucleotides comprising fragments of polynucleotides encoding SEQ ID NO: 2.

The remaining claims are rejected for depending from an indefinite claim.

Applicants' 12/6/06 response at page 9 states "[p]revious claims 68-92 were also rejected under 35 USC 112 second paragraph. This rejection is not applicable to the current claims which do not recite "contained in"..."

Applicants' arguments have been fully considered but are not found persuasive. Contrary to Applicants' statement at pg 9 of the response, new claims 93-97 each recite the phrase "contained in" one or more times. Specifically, it is noted that claim 93 uses the phrase "contained in" in five instances (in lines 3, 4, 5, 7, 10); claim 94 uses the

same phrase in two instances (in lines 2 and 4); claim 95 uses the same phrase in two instances (in lines 2 and 4); and claims 96 and 97 use the same phrase in one instance (each in line 2). It is noted that this rejection would be withdrawn if the claims were amended, for example, to recite "of" in place of each usage of "contained in" (for example, claim 96 would recite "the sequence of SEQ ID NO: 1").

New objections and/or rejections
Claim Objections

Claim 93 is objected to because of the following informalities:

(1) In claim 93, line 3, the word "polypeptide" is misspelled as "polypetide".

Appropriate correction is required.

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 106-108 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

Claim 106, as written, does not sufficiently distinguish over cells that exist naturally because the claims do not particularly point out any non-naturally occurring differences between the claimed products and the naturally occurring products. Specifically, the claim encompasses non-isolated host cells found within an organism that contain the nucleic acid sequences encompassed by the claim. In the absence of the hand of man, the naturally occurring products are considered non-statutory subject matter. See Diamond v. Chakrabarty, 447 U.S. 303, 206 USPQ 193 (1980). Claim 106 should be amended to indicate the hand of the inventor, e.g. by amendment to recite "An isolated cell..." See MPEP 2105.

Claims 107 and 108 depend from claim 106 and are included in the rejection because they each encompass non-isolated host cells.

Claim Rejections - 35 USC § 112, 1st paragraph, new matter

Claims 93-117 are also rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement because the claims contain new matter.

New claim 93 recites, "...wherein high stringency conditions are incubating for 5 minutes in 0.02X SSC at 65 degrees C..." However, the specification as originally filed does contain support for this limitation. The specification at page 16, paragraph [0044] teaches, "An example of highly stringent wash conditions is 15 minutes in 0.1 X SSC at 65°C. An example of stringent was conditions is 15 minutes in 0.2 X SSC buffer at 65°C." There is no conception in the specification of the limitation of "0.02X SSC" as an example of high stringency condition, nor does the concept of the specific limitation flow naturally from the disclosure of the specification. Therefore, the specification as originally filed lacks support for this limitation. Claims 94-117 are included in this rejection because they depend from claim 93.

Claim Rejections - 35 USC § 102a

Claims 93-98 are rejected under 35 U.S.C. 102(a) as being anticipated by NCBI Entrez Nucleotide, database record for Locus AC092214, Version AC092214.2, 15 September 2001, 60 pages. The earliest date to which the instant application claims priority is July 29th, 2002.

The database record for Locus AC092214, version AC092214.2, teaches a DNA sequence of 209,501 base pairs designated "Homo sapiens chromosome 7 clone RP11-811J9, working draft sequence, 3 unordered pieces". An alignment between this DNA sequence and SEQ ID NO: 1 of the instant application indicates that residues 158107-157151 of AC092214.2 are 99.9% similar to instant SEQ ID NO: 1 (an alignment of these sequences is attached to this Office Action as Sequence Alignment #1). The only difference between the two sequences is that residue 930 of SEQ ID NO: 1, is listed as 'Y', which indicates a pyrimidine residue (either a thymine ('T') or cytosine ('C')), and that residue 157236 of AC092214.2 is listed as a thymine reside ('T').

Because a pyrimidine residue indicates a genus that includes either thymine or cytosine, the sequence of AC092214.2 effectively teaches two sequences, one of which comprises a species encompassed by SEQ ID NO: 1. Furthermore, an alignment between the protein encoded by AC092214.2 and SEQ ID NO: 2 (the polypeptide encoded by SEQ ID NO: 1) indicates that residues 158107-157151 of AC092214.2 encode a polypeptide that is 100% identical to instant SEQ ID NO: 2 (an alignment of these sequences is attached to this Office Action as Sequence Alignment #2). The record for AC092214.2 indicates that this sequence is "Homo sapiens chromosome 7 clone RP11-811J9" which meets the limitation of "isolated". Therefore, because AC092214.2 encodes a polypeptide that is 100% identical to instant SEQ ID NO: 2, this isolated DNA sequence anticipates each of instant claims 93-95 and 97. Furthermore, because AC092214.2 comprises one of the two sequences "contained in" SEQ ID NO: 1, AC092214.2 also anticipates claim 96. Finally, the record for AC092214.2 indicates the sequence is from "genomic DNA" (see page 2) and therefore AC092214.2 also anticipates instant claim 98, which depends from claim 93 and encompasses a "genomic sequence".

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 93, 96, 97 and 98 are rejected under 35 U.S.C. 102(b) as being anticipated by NCBI Entrez Nucleotide, database record for Locus AC092214, Version AC092214.1, 29 June 2001, 76 pages. The earliest date to which the instant application claims priority is July 29th, 2002.

This rejection is based on claims 93 and 96-98 encompassing polynucleotides comprising fragments of SEQ ID NO: 1, or comprising fragments of polynucleotides encoding SEQ ID NO: 2. See "Claim Rejections - 35 U.S.C. 112, 2nd Paragraph.

The database record for Locus AC092214, version AC092214.1, teaches a DNA sequence of 266224 base pairs designated "Homo sapiens chromosome UNK clone RP11-811J9, working draft sequence, 6 unordered pieces". An alignment between this DNA sequence and instant SEQ ID NO: 1 indicates an overall 83.9% similarity (an alignment of these sequences is attached to this Office Action as Sequence Alignment #3). Specifically, residues 7960-8802 of AC092214.1 are 97.6% similar to residues 117-957 of instant SEQ ID NO: 1, for an overall 83.9% similarity. Furthermore, an alignment between the protein encoded by AC092214.2 and SEQ ID NO: 2 (the polypeptide encoded by SEQ ID NO: 1) indicates that the encoded proteins share an overall 83.8% similarity (an alignment of these sequences is attached to this Office Action as Sequence Alignment #4). Specifically, residues 7941-8799 of AC092214.1 encode a protein that is 96.5% identical to residues 34-318 of SEQ ID NO: 2, for an overall 83.8% similarity between the two proteins. Furthermore, within the region of identity between the polynucleotide sequence of AC092214.1 and SEQ ID NO: 1 are numerous sequence of two or more polynucleotides that are 100% identical. Therefore, the polynucleotide sequence of AC092214.1 "has the sequence contained in SEQ ID NO: 1" and "encodes a polypeptide comprising the sequence contained in SEQ ID NO: 2", and therefore anticipates claims 93, 96 and 97. Finally, the record for AC092214.1 indicates the sequence is from "genomic DNA" (see page 2) and therefore AC092214.1 also anticipates instant claim 98, which depends from claim 93 and encompasses a "genomic sequence".

Conclusion

No claims are allowable.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Zachary C. Howard whose telephone number is 571-272-2877. The examiner can normally be reached on M-F 9:30 AM - 6:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary B. Nickol can be reached on 571-272-0835. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

zch

[signature]

GARY B. NICKOL, PH.D.
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 1600

10628464-1_vs_AC092214.2

GenCore version 6.2
Copyright (c) 1993 - 2007 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 2, 2007, 15:17:51 ; Search time 21 Seconds
                                          (without alignments)
                                          19.095 Million cell updates/sec

Title: US-10-628-464-1
Perfect score: 956.6
Sequence: 1 atgaatggagaccacatggt..........caaggtgtgggacaccttga 957

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 209501 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   acv2.txt:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                              SUMMARIES
                          %
Result   Query
   No.   Score Match Length DB ID   Description
----
 c 1   956.6 100.0 209501 1   AC092214   ACCESSION:AC092214
    2   26.4   2.8 209501 1   AC092214   ACCESSION:AC092214

                              ALIGNMENTS

RESULT 1
AC092214/c
LOCUS   AC092214   209501 bp   DNA   linear   HTG 15-SEP-2001
DEFINITION   Homo sapiens chromosome 7 clone RP11-811J9, WORKING DRAFT SEQUENCE,
               3 unordered pieces.
ACCESSION   AC092214
VERSION   AC092214.2 GI:15625026
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 209501)
  AUTHORS   Waterston,R.H.
  TITLE   The sequence of Homo sapiens clone
  JOURNAL   Unpublished

```
REFERENCE   2  (bases 1 to 209501)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
  COMMENT   [WARNING] On Nov 15, 2001 this sequence was replaced by
            gi:16931065.
            On Sep 15, 2001 this sequence version replaced gi:14575807.


            ------------------ Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------------- Project Information ----------------
            Center project name: H_NH0811J09
            ----------------------------------------------------
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 1%
            Sequencing vector: plasmid; 99%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 208290 bases at least Q40
            Consensus quality: 208711 bases at least Q30
            Consensus quality: 209110 bases at least Q20
            ----------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    4259: contig of 4259 bp in length
            *    4260    4359: gap of unknown length
            *    4360    9850: contig of 5491 bp in length
            *    9851    9950: gap of unknown length
            *    9951  209501: contig of 199551 bp in length.
FEATURES             Location/Qualifiers
     source          1..209501
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /clone="RP11-811J9"
     misc_feature    1..4259
                     /note="assembly_name:Contig11"
     gap             4260..4359
                     /estimated_length=unknown
     misc_feature    4360..9850
                     /note="assembly_name:Contig12"
     gap             9851..9950
                     /estimated_length=unknown
     misc_feature    9951..209501
                     /note="assembly_name:Contig13"

   Query Match            100.0%;  Score 956.6;  DB 1;  Length 209501;
   Best Local Similarity   99.9%;  Pred. No. 0;
   Matches  956;  Conservative   1;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       1 ATGAATGGAGACCACATGGTTCTAGGATCTTCGGTGACTGACAAGAAGGCCATCATCTTG 60
```

10628464-1_vs_AC092214.2

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   158107  ATGAATGGAGACCACATGGTTCTAGGATCTTCGGTGACTGACAAGAAGGCCATCATCTTG  158048

Qy       61  GTTACCATTTTACTCCTTTTACGCCTGGTAGCAATAGCAGGCAATGGCTTCATCACTGCT  120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   158047  GTTACCATTTTACTCCTTTTACGCCTGGTAGCAATAGCAGGCAATGGCTTCATCACTGCT  157988

Qy      121  GCTCTGGGCGTGGAGTGGGTGCTACGGAGAATGTTGTTGCCTTGTGATAAGTTATTGGTT  180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157987  GCTCTGGGCGTGGAGTGGGTGCTACGGAGAATGTTGTTGCCTTGTGATAAGTTATTGGTT  157928

Qy      181  AGCCTAGGGGCCTCTCGCTTCTGTCTGCAGTCAGTGGTAATGGGTAAGACCATTTATGTT  240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157927  AGCCTAGGGGCCTCTCGCTTCTGTCTGCAGTCAGTGGTAATGGGTAAGACCATTTATGTT  157868

Qy      241  TTCTTGCATCCGATGGCCTTCCCATACAACCCTGTACTGCAGTTTCTAGCTTTCCAGTGG  300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157867  TTCTTGCATCCGATGGCCTTCCCATACAACCCTGTACTGCAGTTTCTAGCTTTCCAGTGG  157808

Qy      301  GACTTCCTGAATGCTGCCACCTTATGGTCCTCTACCTGGCTCAGTGTCTTCTATTGTGTG  360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157807  GACTTCCTGAATGCTGCCACCTTATGGTCCTCTACCTGGCTCAGTGTCTTCTATTGTGTG  157748

Qy      361  AAAATTGCTACCTTCACCCACCCTGTCTTCTTCTGGCTAAAGCACAAGTTGTCTGGGTGG  420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157747  AAAATTGCTACCTTCACCCACCCTGTCTTCTTCTGGCTAAAGCACAAGTTGTCTGGGTGG  157688

Qy      421  CTACCATGGATGCTCTTCAGCTCTGTAGGGCTCTCCAGCTTCACCACCATTCTATTTTTC  480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157687  CTACCATGGATGCTCTTCAGCTCTGTAGGGCTCTCCAGCTTCACCACCATTCTATTTTTC  157628

Qy      481  ATAGGCAACCACAGAATGTATCAGAACTATTTAAGGAACCATCTACAACCTTGGAATGTC  540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157627  ATAGGCAACCACAGAATGTATCAGAACTATTTAAGGAACCATCTACAACCTTGGAATGTC  157568

Qy      541  ACTGGCGATAGCATACGGAGCTACTGTGAGAAATTCTATCTCTTCCCTCTAAAAATGATT  600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157567  ACTGGCGATAGCATACGGAGCTACTGTGAGAAATTCTATCTCTTCCCTCTAAAAATGATT  157508

Qy      601  ACTTGGACAATGCCCACTGCTGTCTTTTTCATTTGCATGATTTTGCTCATCACATCTCTG  660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157507  ACTTGGACAATGCCCACTGCTGTCTTTTTCATTTGCATGATTTTGCTCATCACATCTCTG  157448

Qy      661  GGAAGACACAGGAAGAAGGCTCTCCTTACAACCTCAGGATTCCGAGAGCCCAGTGTGCAG  720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157447  GGAAGACACAGGAAGAAGGCTCTCCTTACAACCTCAGGATTCCGAGAGCCCAGTGTGCAG  157388

Qy      721  GCACACATAAAGGCTCTGCTGGCTCTCCTCTCTTTTGCCATGCTCTTCATCTCATATTTC  780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157387  GCACACATAAAGGCTCTGCTGGCTCTCCTCTCTTTTGCCATGCTCTTCATCTCATATTTC  157328

Qy      781  CTGTCACTGGTGTTCAGTGCTGCAGGTATTTTTCCACCTCTGGACTTTAAATTCTGGGTG  840
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157327  CTGTCACTGGTGTTCAGTGCTGCAGGTATTTTTCCACCTCTGGACTTTAAATTCTGGGTG  157268

Qy      841  TGGGAGTCAGTGATTTATCTGTGTGCAGCAGTTCACCCCATCATTCTGCTCTTCAGCAAC  900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157267  TGGGAGTCAGTGATTTATCTGTGTGCAGCAGTTCACCCCATCATTCTGCTCTTCAGCAAC  157208

Qy      901  TGCAGGCTGAGAGCTGTGCTGAAGAGTCGYCGTTCCTCAAGGTGTGGGACACCTTGA  957
             ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db   157207  TGCAGGCTGAGAGCTGTGCTGAAGAGTCGTCGTTCCTCAAGGTGTGGGACACCTTGA  157151
```

ac092214.2

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 25, 2007, 15:53:17 ; Search time 61 Seconds
                                             (without alignments)
                                             3.276 Million cell updates/sec

Title: US-10-628-464-2
Perfect score: 1678
Sequence: 1 MNGDHMVLGSSVTDKKAIIL..........SNCRLRAVLKSRRSSRCGTP 318

Scoring table: BLOSUM62
                Xgapop 10.0 , Xgapext 0.5
                Ygapop 10.0 , Ygapext 0.5
                Fgapop 6.0 , Fgapext 7.0
                Delop   6.0 , Delext   7.0

Searched: 1 seqs, 209501 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 500 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=US10628464.pep -DB=acv2.txt -SUFFIX=pto
-OUT=ac092214.2.res -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1
-END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY
-NEG_SCORES=0 -LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   acv2.txt:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result   Query
    No.   Score   Match Length DB ID   Description
        ----
  c 1   1678   100.0 209501 1   AC092214   ACCESSION:AC092214
     2   80   4.8 209501 1   AC092214   ACCESSION:AC092214

                              ALIGNMENTS

RESULT 1
AC092214/c
LOCUS   AC092214   209501 bp   DNA   linear   HTG 15-SEP-2001
DEFINITION Homo sapiens chromosome 7 clone RP11-811J9, WORKING DRAFT SEQUENCE,
            3 unordered pieces.
                              Page 1

```
ACCESSION   AC092214
VERSION     AC092214.2  GI:15625026
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 209501)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 209501)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     [WARNING] On Nov 15, 2001 this sequence was replaced by
            gi:16931065.
            On Sep 15, 2001 this sequence version replaced gi:14575807.

            -------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------------- Project Information ----------------
            Center project name: H_NH0811J09
            ------------------------------------------------------
            -------------- Summary Statistics --------------
            Sequencing vector: M13; 1%
            Sequencing vector: plasmid; 99%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 208290 bases at least Q40
            Consensus quality: 208711 bases at least Q30
            Consensus quality: 209110 bases at least Q20
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     4259: contig of 4259 bp in length
            *    4260     4359: gap of unknown length
            *    4360     9850: contig of 5491 bp in length
            *    9851     9950: gap of unknown length
            *    9951   209501: contig of 199551 bp in length.
FEATURES            Location/Qualifiers
     source         1..209501
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="7"
                    /clone="RP11-811J9"
     misc_feature   1..4259
                    /note="assembly_name:Contig11"
     gap            4260..4359
                    /estimated_length=unknown
     misc_feature   4360..9850
```

ac092214.2
/note="assembly_name:Contig12"
gap 9851..9950
/estimated_length=unknown
misc_feature 9951..209501
/note="assembly_name:Contig13"

Alignment Scores:
Pred. No.:   0   Length: 209501
Score: 1678.00   Matches: 318
Percent Similarity: 100.0%   Conservative: 0
Best Local Similarity: 100.0%   Mismatches: 0
Query Match: 100.0%   Indels: 0
DB: 1   Gaps: 0

US-10-628-464-2 (1-318) x AC092214 (1-209501)

```
Qy       1 MetAsnGlyAspHisMetValLeuGlySerSerValThrAspLysLysAlaIleIleLeu 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  158107 ATGAATGGAGACCACATGGTTCTAGGATCTTCGGTGACTGACAAGAAGGCCATCATCTTG 158048

Qy      21 ValThrIleLeuLeuLeuLeuArgLeuValAlaIleAlaGlyAsnGlyPheIleThrAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  158047 GTTACCATTTTACTCCTTTTACGCCTGGTAGCAATAGCAGGCAATGGCTTCATCACTGCT 157988

Qy      41 AlaLeuGlyValGluTrpValLeuArgArgMetLeuLeuProCysAspLysLeuLeuVal 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157987 GCTCTGGGCGTGGAGTGGGTGCTACGGAGAATGTTGTTGCCTTGTGATAAGTTATTGGTT 157928

Qy      61 SerLeuGlyAlaSerArgPheCysLeuGlnSerValValMetGlyLysThrIleTyrVal 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157927 AGCCTAGGGGCCTCTCGCTTCTGTCTGCAGTCAGTGGTAATGGGTAAGACCATTTATGTT 157868

Qy      81 PheLeuHisProMetAlaPheProTyrAsnProValLeuGlnPheLeuAlaPheGlnTrp 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157867 TTCTTGCATCCGATGGCCTTCCCATACAACCCTGTACTGCAGTTTCTAGCTTTCCAGTGG 157808

Qy     101 AspPheLeuAsnAlaAlaThrLeuTrpSerSerThrTrpLeuSerValPheTyrCysVal 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157807 GACTTCCTGAATGCTGCCACCTTATGGTCCTCTACCTGGCTCAGTGTCTTCTATTGTGTG 157748

Qy     121 LysIleAlaThrPheThrHisProValPhePheTrpLeuLysHisLysLeuSerGlyTrp 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157747 AAAATTGCTACCTTCACCCACCCTGTCTTCTTCTGGCTAAAGCACAAGTTGTCTGGGTGG 157688

Qy     141 LeuProTrpMetLeuPheSerSerValGlyLeuSerSerPheThrThrIleLeuPhePhe 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157687 CTACCATGGATGCTCTTCAGCTCTGTAGGGCTCTCCAGCTTCACCACCATTCTATTTTTC 157628

Qy     161 IleGlyAsnHisArgMetTyrGlnAsnTyrLeuArgAsnHisLeuGlnProTrpAsnVal 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157627 ATAGGCAACCACAGAATGTATCAGAACTATTTAAGGAACCATCTACAACCTTGGAATGTC 157568

Qy     181 ThrGlyAspSerIleArgSerTyrCysGluLysPheTyrLeuPheProLeuLysMetIle 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157567 ACTGGCGATAGCATACGGAGCTACTGTGAGAAATTCTATCTCTTCCCTCTAAAAATGATT 157508

Qy     201 ThrTrpThrMetProThrAlaValPhePheIleCysMetIleLeuLeuIleThrSerLeu 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157507 ACTTGGACAATGCCCACTGCTGTCTTTTTCATTTGCATGATTTTGCTCATCACATCTCTG 157448

Qy     221 GlyArgHisArgLysLysAlaLeuLeuThrThrSerGlyPheArgGluProSerValGln 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157447 GGAAGACACAGGAAGAAGGCTCTCCTTACAACCTCAGGATTCCGAGAGCCCAGTGTGCAG 157388
```

Page 3

```
Qy       241 AlaHisIleLysAlaLeuLeuAlaLeuLeuSerPheAlaMetLeuPheIleSerTyrPhe 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    157387 GCACACATAAAGGCTCTGCTGGCTCTCCTCTCTTTGCCATGCTCTTCATCTCATATTTC 157328

Qy       261 LeuSerLeuValPheSerAlaAlaGlyIlePheProProLeuAspPheLysPheTrpVal 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    157327 CTGTCACTGGTGTTCAGTGCTGCAGGTATTTTTCCACCTCTGGACTTTAAATTCTGGGTG 157268

Qy       281 TrpGluSerValIleTyrLeuCysAlaAlaValHisProIleIleLeuLeuPheSerAsn 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    157267 TGGGAGTCAGTGATTTATCTGTGTGCAGCAGTTCACCCCATCATTCTGCTCTTCAGCAAC 157208

Qy       301 CysArgLeuArgAlaValLeuLysSerArgArgSerSerArgCysGlyThrPro 318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||·||
Db    157207 TGCAGGCTGAGAGCTGTGCTGAAGAGTCGTCGTTCCTCAAGGTGTGGGACACCT 157154
```

RESULT 2
AC092214
```
LOCUS       AC092214               209501 bp    DNA     linear   HTG 15-SEP-2001
DEFINITION  Homo sapiens chromosome 7 clone RP11-811J9, WORKING DRAFT SEQUENCE,
            3 unordered pieces.
ACCESSION   AC092214
VERSION     AC092214.2  GI:15625026
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 209501)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 209501)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     [WARNING] On Nov 15, 2001 this sequence was replaced by
            gi:16931065.
            On Sep 15, 2001 this sequence version replaced gi:14575807.

            ------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ------------------ Project Information ----------------
            Center project name: H_NH0811J09
            ----------------------------------------------------------
            -------------- Summary Statistics  ----------------
            Sequencing vector: M13; 1%
            Sequencing vector: plasmid; 99%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 208290 bases at least Q40
            Consensus quality: 208711 bases at least Q30
            Consensus quality: 209110 bases at least Q20
            ---------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
```

SEQUENCE ALIGNMENT #3

10628464-1_vs_AC092214.1

GenCore version 6.2
Copyright (c) 1993 - 2007 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 2, 2007, 15:16:21 ; Search time 26 Seconds
                                    (without alignments)
                                 19.598 Million cell updates/sec

Title: US-10-628-464-1
Perfect score: 956.6
Sequence: 1 atgaatggagaccacatggt..........caaggtgtgggacaccttga 957

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 266224 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   acv1.txt:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Length	DB	ID	Description
1	802.4	83.9	266224	1	AC092214	ACCESSION:AC092214
c 2	252.8	26.4	266224	1	AC092214	ACCESSION:AC092214

ALIGNMENTS

RESULT 1
AC092214
LOCUS   AC092214   266224 bp   DNA   linear   HTG 29-JUN-2001
DEFINITION   Homo sapiens chromosome UNK clone RP11-811J9, WORKING DRAFT
             SEQUENCE, 6 unordered pieces.
ACCESSION   AC092214
VERSION   AC092214.1 GI:14575807
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 266224)
  AUTHORS   Waterston,R.H.
  TITLE   The sequence of Homo sapiens clone
  JOURNAL   Unpublished

```
REFERENCE   2  (bases 1 to 266224)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
  COMMENT   [WARNING] On Sep 15, 2001 this sequence was replaced by
            gi:15625026.

            ------------------ Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------------- Project Information ----------------
            Center project name: H_NH0811J09
            ----------------------------------------------------
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 1%
            Sequencing vector: plasmid; 98%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 99% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 203494 bases at least Q40
            Consensus quality: 204566 bases at least Q30
            Consensus quality: 264940 bases at least Q20
            Insert size: 210000; agarose-fp
            Insert size: 265724; sum-of-contigs
            Quality coverage: 7.96 in Q20 bases; agarose-fp
            Quality coverage: 6.30 in Q20 bases; sum-of-contigs
            ----------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    1883: contig of 1883 bp in length
            *     1884    1983: gap of unknown length
            *     1984    3064: contig of 1081 bp in length
            *     3065    3164: gap of unknown length
            *     3165    7823: contig of 4659 bp in length
            *     7824    7923: gap of unknown length
            *     7924   14371: contig of 6448 bp in length
            *    14372   14471: gap of unknown length
            *    14472   65512: contig of 51041 bp in length
            *    65513   65612: gap of unknown length
            *    65613  266224: contig of 200612 bp in length.
FEATURES             Location/Qualifiers
     source          1..266224
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="UNK"
                     /clone="RP11-811J9"
     misc_feature    1..1883
                     /note="assembly_name:Contig11"
     gap             1884..1983
                     /estimated_length=unknown
     misc_feature    1984..3064
                     /note="assembly_name:Contig12"
     gap             3065..3164
```

```
                                    /estimated_length=unknown
         misc_feature          3165..7823
                                    /note="assembly_name:Contig13"
         gap                   7824..7923
                                    /estimated_length=unknown
         misc_feature          7924..14371
                                    /note="assembly_name:Contig14"
         gap                   14372..14471
                                    /estimated_length=unknown
         misc_feature          14472..65512
                                    /note="assembly_name:Contig15"
         gap                   65513..65612
                                    /estimated_length=unknown
         misc_feature          65613..266224
                                    /note="assembly_name:Contig16"
```

```
   Query Match              83.9%;   Score 802.4;   DB 1;   Length 266224;
   Best Local Similarity    97.6%;   Pred. No. 0;
   Matches  823;  Conservative   1;  Mismatches   17;  Indels   2;  Gaps    1;
```

```
Qy      117 TGCTGCTCTGGGCGTGGAGTGGGTG--CTACGGAGAATGTTGTTGCCTTGTGATAAGTTA 174
            | |||||| ||||| ||||| | |      | ||||||||||||||||| |||||||||||
Db     7960 TTCTGCTGTGGGAGTGGAGGGGGTGGCTACGGAGGAATGTTGTTGCCTGGTGATAAGTTA 8019

Qy      175 TTGGTTAGCCTAGGGGCCTCTCGCTTCTGTCTGCAGTCAGTGGTAATGGGTAAGACCATT 234
            ||||||||| ||||||||||||||||||| ||||||||||||||||||   |||||||||||
Db     8020 TTGGTTAGCTTAGGGGCCTCTCGCTTCTTTCTGCAGTCAGTGGTAAGGGGTAAGACCATT 8079

Qy      235 TATGTTTTCTTGCATCCGATGGCCTTCCCATACAACCCTGTACTGCAGTTTCTAGCTTTC 294
            |||||||| |||||||||||||||||||||||||||| |||||||||||||||||||||||
Db     8080 TATGTTTTGTTGCATCCGATGGCCTTCCCATACAACCGTGTACTGCAGTTTCTAGCTTTC 8139

Qy      295 CAGTGGGACTTCCTGAATGCTGCCACCTTATGGTCCTCTACCTGGCTCAGTGTCTTCTAT 354
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8140 CAGTGGGACTTCCTGAATGCTGCCACCTTATGGTCCTCTACCTGGCTCAGTGTCTTCTAT 8199

Qy      355 TGTGTGAAAATTGCTACCTTCACCCACCCTGTCTTCTTCTGGCTAAAGCACAAGTTGTCT 414
            |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db     8200 TGTGTGAAAATTGCTACCTTCACCCACCCTGTCTTCTTCTGGGTAAAGCACAAGTTGTCT 8259

Qy      415 GGGTGGCTACCATGGATGCTCTTCAGCTCTGTAGGGCTCTCCAGCTTCACCACCATTCTA 474
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8260 GGGTGGCTACCATGGATGCTCTTCAGCTCTGTAGGGCTCTCCAGCTTCACCACCATTCTA 8319

Qy      475 TTTTTCATAGGCAACCACAGAATGTATCAGAACTATTTAAGGAACCATCTACAACCTTGG 534
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8320 TTTTTCATAGGCAACCACAGAATGTATCAGAACTATTTAAGGAACCATCTACAACCTTGG 8379

Qy      535 AATGTCACTGGCGATAGCATACGGAGCTACTGTGAGAAATTCTATCTCTTCCCTCTAAAA 594
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8380 AATGTCACTGGCGATAGCATACGGAGCTACTGTGAGAAATTCTATCTCTTCCCTCTAAAA 8439

Qy      595 ATGATTACTTGGACAATGCCCACTGCTGTCTTTTTCATTTGCATGATTTTGCTCATCACA 654
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8440 ATGATTACTTGGACAATGCCCACTGCTGTCTTTTTCATTTGCATGATTTTGCTCATCACA 8499

Qy      655 TCTCTGGGAAGACACAGGAAGAAGGCTCTCCTTACAACCTCAGGATTCCGAGAGCCCAGT 714
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8500 TCTCTGGGAAGACACAGGAAGAAGGCTCTCCTTACAACCTCAGGATTCCGAGAGCCCAGT 8559

Qy      715 GTGCAGGCACACATAAAGGCTCTGCTGGCTCTCCTCTCTTTTGCCATGCTCTTCATCTCA 774
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8560 GTGCAGGCACACATAAAGGCTCTGCTGGCTCTCCTCTCTTTTGCCATGCTCTTCATCTCA 8619
```

```
Qy    775 TATTTCCTGTCACTGGTGTTCAGTGCTGCAGGTATTTTTCCACCTCTGGACTTTAAATTC 834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8620 TATTTCCTGTCACTGGTGTTCAGTGCTGCAGGTATTTTTCCACCTCTGGACTTTAAATTC 8679

Qy    835 TGGGTGTGGGAGTCAGTGATTTATCTGTGTGCAGCAGTTCACCCCATCATTCTGCTCTTC 894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8680 TGGGTGTGGGAGTCAGTGATTTATCTGTGTGCAGCAGTTCACCCCATCATTCTGCTCTTC 8739

Qy    895 AGCAACTGCAGGCTGAGAGCTGTGCTGAAGAGTCGYCGTTCCTCAAGGTGTGGGACACCT 954
          |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db   8740 AGCAACTGCAGGCTGAGAGCTGTGCTGAAGAGTCGTCGTTCCTCAAGGTGTGGGACACCT 8799

Qy    955 TGA 957
          |||
Db   8800 TGA 8802
```

RESULT 2
AC092214/c
```
LOCUS       AC092214              266224 bp    DNA     linear   HTG 29-JUN-2001
DEFINITION  Homo sapiens chromosome UNK clone RP11-811J9, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
ACCESSION   AC092214
VERSION     AC092214.1  GI:14575807
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 266224)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 266224)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     [WARNING] On Sep 15, 2001 this sequence was replaced by
            gi:15625026.

            ------------------- Genome Center -------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ------------------- Project Information -------------------
            Center project name: H_NH0811J09
            ------------------- Summary Statistics -------------------
            Sequencing vector: M13; 1%
            Sequencing vector: plasmid; 98%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 99% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 203494 bases at least Q40
            Consensus quality: 204566 bases at least Q30
            Consensus quality: 264940 bases at least Q20
            Insert size: 210000; agarose-fp
            Insert size: 265724; sum-of-contigs
            Quality coverage: 7.96 in Q20 bases; agarose-fp
            Quality coverage: 6.30 in Q20 bases; sum-of-contigs
```

ac092214.1

GenCore version 5.1.9
Copyright (c) 1993 - 2007 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 25, 2007, 15:37:56 ; Search time 78 Seconds
                                            (without alignments)
                                            3.256 Million cell updates/sec

Title: US-10-628-464-2
Perfect score: 1678
Sequence: 1 MNGDHMVLGSSVTDKKAIIL..........SNCRLRAVLKSRRSSRCGTP 318

Scoring table: BLOSUM62
                Xgapop 10.0 , Xgapext 0.5
                Ygapop 10.0 , Ygapext 0.5
                Fgapop 6.0 , Fgapext 7.0
                Delop   6.0 , Delext   7.0

Searched: 1 seqs, 266224 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 500 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=US10628464.pep -DB=acv1.txt -SUFFIX=pto
-OUT=ac092214.1.res -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1
-END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY
-NEG_SCORES=0 -LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   acv1.txt:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                                SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1407	83.8	266224	1	AC092214	ACCESSION:AC092214
c 2	609	36.3	266224	1	AC092214	ACCESSION:AC092214

                                ALIGNMENTS

RESULT 1
AC092214
LOCUS   AC092214   266224 bp   DNA   linear   HTG 29-JUN-2001
DEFINITION Homo sapiens chromosome UNK clone RP11-811J9, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
                                Page 1

```
ACCESSION   AC092214
VERSION     AC092214.1  GI:14575807
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 266224)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 266224)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     [WARNING] On Sep 15, 2001 this sequence was replaced by
            gi:15625026.

            ------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------------- Project Information ----------------
            Center project name: H_NH0811J09
            ----------------------------------------------------
            -------------- Summary Statistics --------------
            Sequencing vector: M13; 1%
            Sequencing vector: plasmid; 98%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 99% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 203494 bases at least Q40
            Consensus quality: 204566 bases at least Q30
            Consensus quality: 264940 bases at least Q20
            Insert size: 210000; agarose-fp
            Insert size: 265724; sum-of-contigs
            Quality coverage: 7.96 in Q20 bases; agarose-fp
            Quality coverage: 6.30 in Q20 bases; sum-of-contigs
            ----------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1883: contig of 1883 bp in length
            *     1884     1983: gap of unknown length
            *     1984     3064: contig of 1081 bp in length
            *     3065     3164: gap of unknown length
            *     3165     7823: contig of 4659 bp in length
            *     7824     7923: gap of unknown length
            *     7924    14371: contig of 6448 bp in length
            *    14372    14471: gap of unknown length
            *    14472    65512: contig of 51041 bp in length
            *    65513    65612: gap of unknown length
            *    65613   266224: contig of 200612 bp in length.
FEATURES             Location/Qualifiers
     source          1..266224
                     /organism="Homo sapiens"
```

```
                              ac092214.1
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="UNK"
                    /clone="RP11-811J9"
        misc_feature    1..1883
                    /note="assembly_name:Contig11"
        gap             1884..1983
                    /estimated_length=unknown
        misc_feature    1984..3064
                    /note="assembly_name:Contig12"
        gap             3065..3164
                    /estimated_length=unknown
        misc_feature    3165..7823
                    /note="assembly_name:Contig13"
        gap             7824..7923
                    /estimated_length=unknown
        misc_feature    7924..14371
                    /note="assembly_name:Contig14"
        gap             14372..14471
                    /estimated_length=unknown
        misc_feature    14472..65512
                    /note="assembly_name:Contig15"
        gap             65513..65612
                    /estimated_length=unknown
        misc_feature    65613..266224
                    /note="assembly_name:Contig16"
```

Alignment Scores:

Pred. No.:	0	Length:	266224
Score:	1407.00	Matches:	276
Percent Similarity:	96.5%	Conservative:	3
Best Local Similarity:	95.5%	Mismatches:	6
Query Match:	83.8%	Indels:	4
DB:	1	Gaps:	0

US-10-628-464-2 (1-318) x AC092214 (1-266224)

```
Qy      34 GlyAsnGlyPhe-IleThr-AlaAlaLeuGlyValGluTrpVal-LeuArgArg-MetLe  52
           ||||||||||||| ||||||| :::|||:::|||||||||| ||| |||||||||| |||||
Db    7941 GGCAATGGCTTCAATCACTTTCTGCTGTGGGAGTGGAGGGGGTGGCTACGGAGGAATGTT  8000

Qy      52 uLeuProCysAspLysLeuLeuValSerLeuGlyAlaSerArgPheCysLeuGlnSerVa  72
           |||||||  |||||||||||||||||||||||||||   ||||||||| |||||||||||
Db    8001 GTTGCCTGGTGATAAGTTATTGGTTAGCTTAGGGGCCTCTCGCTTCTTTCTGCAGTCAGT  8060

Qy      72 lValMetGlyLysThrIleTyrValPheLeuHisProMetAlaPheProTyrAsnProVa  92
           ||||   |||||||||||||||||||   ||||||||||||||||||||||||||   ||
Db    8061 GGTAAGGGGTAAGACCATTTATGTTTTGTTGCATCCGATGGCCTTCCCATACAACCGTGT  8120

Qy      92 lLeuGlnPheLeuAlaPheGlnTrpAspPheLeuAsnAlaAlaThrLeuTrpSerSerTh  112
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8121 ACTGCAGTTTCTAGCTTTCCAGTGGGACTTCCTGAATGCTGCCACCTTATGGTCCTCTAC  8180

Qy     112 rTrpLeuSerValPheTyrCysValLysIleAlaThrPheThrHisProValPhePheTr  132
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8181 CTGGCTCAGTGTCTTCTATTGTGTGAAAATTGCTACCTTCACCCACCCTGTCTTCTTCTG  8240

Qy     132 pLeuLysHisLysLeuSerGlyTrpLeuProTrpMetLeuPheSerSerValGlyLeuSe  152
           |:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8241 GGTAAAGCACAAGTTGTCTGGGTGGCTACCATGGATGCTCTTCAGCTCTGTAGGGCTCTC  8300

Qy     152 rSerPheThrThrIleLeuPhePheIleGlyAsnHisArgMetTyrGlnAsnTyrLeuAr  172
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Page 3

```
Db    8301 CAGCTTCACCACCATTCTATTTTTCATAGGCAACCACAGAATGTATCAGAACTATTTAAG 8360

Qy     172 gAsnHisLeuGlnProTrpAsnValThrGlyAspSerIleArgSerTyrCysGluLysPh 192
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8361 GAACCATCTACAACCTTGGAATGTCACTGGCGATAGCATACGGAGCTACTGTGAGAAATT 8420

Qy     192 eTyrLeuPheProLeuLysMetIleThrTrpThrMetProThrAlaValPhePheIleCy 212
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8421 CTATCTCTTCCCTCTAAAAATGATTACTTGGACAATGCCCACTGCTGTCTTTTTCATTTG 8480

Qy     212 sMetIleLeuLeuIleThrSerLeuGlyArgHisArgLysLysAlaLeuLeuThrThrSe 232
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8481 CATGATTTTGCTCATCACATCTCTGGGAAGACACAGGAAGAAGGCTCTCCTTACAACCTC 8540

Qy     232 rGlyPheArgGluProSerValGlnAlaHisIleLysAlaLeuLeuAlaLeuLeuSerPh 252
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8541 AGGATTCCGAGAGCCCAGTGTGCAGGCACACATAAAGGCTCTGCTGGCTCTCCTCTCTTT 8600

Qy     252 eAlaMetLeuPheIleSerTyrPheLeuSerLeuValPheSerAlaAlaGlyIlePhePr 272
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8601 TGCCATGCTCTTCATCTCATATTTCCTGTCACTGGTGTTCAGTGCTGCAGGTATTTTTCC 8660

Qy     272 oProLeuAspPheLysPheTrpValTrpGluSerValIleTyrLeuCysAlaAlaValHi 292
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8661 ACCTCTGGACTTTAAATTCTGGGTGTGGGAGTCAGTGATTTATCTGTGTGCAGCAGTTCA 8720

Qy     292 sProIleIleLeuLeuPheSerAsnCysArgLeuArgAlaValLeuLysSerArgArgSe 312
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8721 CCCCATCATTCTGCTCTTCAGCAACTGCAGGCTGAGAGCTGTGCTGAAGAGTCGTCGTTC 8780

Qy     312 rSerArgCysGlyThrPro 318
            |||||||||||||||||||
Db    8781 CTCAAGGTGTGGGACACCT 8799
```

RESULT 2
AC092214/c
```
LOCUS       AC092214              266224 bp    DNA     linear   HTG 29-JUN-2001
DEFINITION  Homo sapiens chromosome UNK clone RP11-811J9, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
ACCESSION   AC092214
VERSION     AC092214.1  GI:14575807
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 266224)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 266224)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     [WARNING] On Sep 15, 2001 this sequence was replaced by
            gi:15625026.

            ------------------- Genome Center -------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
```